ERGONOMICS, 2002, VOL. 45, NO. 10, 682–698

Taylor & Francis
Taylor & Francis Group

Effects of hot and cold temperature exposure on performance: a meta-analytic review

JUNE J. PILCHER†*, ERIC NADLER‡ and CAROLINE BUSCH§

†Department of Psychology, Clemson University, Clemson, SC 29634, USA

‡John A. Volpe National Transportation Systems Center, Cambridge, MA 02142, USA

§Supporting Science and Technology, US Army Soldier Center, Natrick, MA 01760, USA

Keywords: Temperature; Performance; Environmental conditions; Meta-analysis.

A meta-analysis to mathematically summarize the effect of hot and cold temperature exposure on performance was completed. The results from 515 effect sizes calculated from 22 original studies suggest that hot and cold temperatures negatively impact performance on a wide range of cognitive-related tasks. More specifically, hot temperatures of 90°F (32.22°C) Web Bulb Globe Temperature Index or above and cold temperatures of 50°F (10°C) or less resulted in the greatest decrement in performance in comparison to neutral temperature conditions (14.88% decrement and 13.91% decrement, respectively). Furthermore, the duration of exposure to the experimental temperature, the duration of exposure to the experimental temperature prior to the task onset, the type of task and the duration of the task had differential effects on performance. The current results indicate that hot and cold temperature exposure have a negative impact on performance and that other variables (e.g., length of exposure to the temperature or task duration) may modify this relationship.

1. Introduction

Adjusting to and working under hot or cold temperatures has long been a challenge for people living under immoderate weather conditions. In spite of the ability in industrialized societies to control indoor temperatures, a similar challenge continues for many people who are regularly exposed to extreme temperatures while working. Concern about the effects of temperature on work-related performance has prompted the National Institute for Occupational Safety and Health (NIOSH) to attempt to establish upper limits for occupational exposure (1972). However, in their revised criteria of 1986, NIOSH did not include upper limits for heat exposure under the assumption that there was not a well-established relationship between performance and heat (NIOSH, 1980).

Although a number of governing bodies have attempted to use the available literature to establish guidelines for on-the-job extreme temperature conditions, they

Ergonomics ISSN 0014-0139 print/ISSN 1366-5847 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/00140130210158419

have met with limited success. This has been largely due to the conflicting nature of the primary studies examining the effect of temperature extremes on performance. For example, some studies have reported little, if any, performance loss under extreme temperature exposure (Chiles 1958, Pepler 1959, Colquhoun 1969, Grether, *et al.* 1971, Ramsey 1975, Ramsey and Pai 1975, Lewis, *et al.* 1983), whereas other studies have reported performance decrements (Mackworth 1947, Fraser 1957, Pepler 1960, Bell, *et al.* 1964, Azer, *et al.* 1972, Fine and Kobrick 1978).

Perhaps not surprisingly, narrative summaries have also had limited success with drawing conclusions from the available data on the effects of temperature extremes on performance (Ramsey 1995). Although there have been numerous narrative attempts to summarize the data (Bell and Provins 1962, Grether 1973, Ramsey and Morrissey 1978, Bell 1981, Kobrick and Fine 1983, Ramsey 1983, Kobrick and Johnson 1991, Ramsey and Kwon 1992), a clear picture of the results has yet to be presented. One possible reason for the confusion in the original data and summaries is the variety of experimental conditions that are used across different studies. For example, the specific type of task, the severity of the temperature exposure, and the duration of the temperature exposure may have differential effects on performance under extreme temperature conditions (Wilkinson 1969, Hancock 1984).

A quantitative approach to summarizing the data would allow an examination of the effects of temperature extremes on performance and an assessment of the effects of variables that may differentially impact performance. A type of quantitative summary examining the relationships among type of task, degree of temperature exposure, and duration of temperature exposure was recently completed (Ramsey and Kwon 1992, Ramsey 1995). In these reviews, figures were used to plot the intersection points of the parameters of interest for each study. Although this approach provided a visual means of summarizing across the studies, it did not permit strong quantitative conclusions.

A meta-analysis (Hunter, *et al.* 1982, Hunter and Schmidt 1990) of the data on temperature exposure and performance should provide much more precise quantitative results than any of the previous summary studies. The meta-analytic technique has been successfully used to quantitatively summarize research studies in a variety of other fields, such as clinical psychology, educational psychology, industrial-organizational psychology, physiological psychology, and social psychology (Goyder and McCutcheon 1995, Huffcutt, *et al.* 1996, Hyman, *et al.* 1989, Klawansky, *et al.* 1995, Paus, *et al.* 1998, Peers and M 1994, Pilcher, *et al.* 2000, Svartberg and Stiles 1991). Meta-analytic reviews, because of their mathematical nature, tend to be objective and consistent. Furthermore, they have several statistical advantages over the more common narrative review. A potential problem with every study that uses a sample of a larger population is that the sample chosen may not actually match the population of interest (i.e., a sampling error). Because it mathematically averages across studies, a meta-analysis actually minimizes the influence of sampling error. Moreover, some original studies may be based on a relatively small sample size, thus creating problems with low power. A meta-analysis avoids this potential problem by not performing significance testing at the individual study level. Effectively, in a meta-analysis, all individual samples are combined into one large sample, which should be largely representative of the general population of interest. An additional advantage of the meta-analytic technique is that it allows an easier definition of the variables (called moderator variables) that may affect the dependent variable

of interest. For example, the length of time that each participant is exposed to an extreme temperature condition is a potential moderator variable.

The purpose of the current study was to use the meta-analytic technique to provide a comprehensive, quantitative analysis of the effects of temperature exposure on performance. To better quantify how temperature exposure may influence performance, the severity of temperature exposure, duration of the experimental session, duration of temperature exposure prior to task onset, type of task, and task duration were defined as potential moderator variables. Because of the conflicting nature of the original studies and the previous narrative reviews, it was difficult to make predictions on the outcome of the meta-analysis. In general, more extreme temperatures were expected to result in a greater performance decrement than moderate temperatures. Predictions on the potential effects of the other moderator variables on performance were not possible.

2. Method

2.1. *Location of study data*

A thorough data search was completed for the current meta-analysis. The American Psychological Association's PsychInfo database was queried for keywords 'thermal', 'temperature', 'hot', 'cold', and 'heat'. The names of the authors that were retrieved from PsychInfo were also submitted as keyword queries. In addition, recent volumes of the journals *Ergonomics* and *Human Factors* were reviewed for relevant articles published too recently to appear in the database. The data search identified 527 articles, reports, and dissertations published between 1922 and 1997. Of these, 226 primary studies examining the effect of environmental temperature conditions on performance were identified. These studies represented a wide range of experimental conditions including many different types of dependent measures, hot and cold temperature exposure, cold water exposure, partial body exposure to temperature conditions, temperature exposure under laboratory conditions, and temperature exposure under field conditions.

2.2. *Decision rules*

The purpose of completing a meta-analysis is to combine the mathematical difference between experimental and control groups across primary studies. To ensure that combining across studies results in meaningful data, decisions must be made in advance for selecting studies for inclusion in the meta-analysis. The following criteria were used for selecting primary studies for inclusion in the current meta-analysis. First, each study had to report hot or cold environmental temperature exposure as an experimental condition. Enough information had to be provided about hot environmental temperature conditions to calculate a Web Bulb Globe Temperature (WBGT) index (described below), if one was not provided. Air temperature had to be provided for cold environmental temperature conditions. Cold water exposure studies and studies where temperature exposure was brought about by clothing or head gear were excluded from the analysis. Second, the neutral temperatures used in the primary studies had to meet the criteria shown in the top portion of table 1. Because the experimental results are compared with the neutral temperature results, it was important to control for this factor by limiting the neutral temperature range that would be allowed in the meta-analysis. The neutral temperature ranges were chosen after reviewing the primary studies to determine what temperatures were

Table 1. Coding characteristics.

Neutral Temperature Ranges
A. 60–69.9°F (15.56–21.06°C) WBGT: hot experimental temperature conditions
B. 65–75°F (18.33–23.89°C): cold experimental temperature conditions
Type of Temperature Exposure
A. Hot (⩾70°F [21.11°C] WBGT)
1. Hot1 (70–79.9°F [21.11–26.61°C] WBGT)
2. Hot2 (80–89.9°F [26.67–32.17°C] WBGT)
3. Hot3 (⩾90°F [32.22°C] WBGT)
B. Cold (<65°F [18.3°C])
1. Cold1 (50–64.9°F [10–18.28°C])
2. Cold2 (<50°F [10°C])
Duration of Experimental Session
A. Short (<120 mins)
B. Long (⩾120 mins)
Duration of Pre-task Temperature Exposure
A. None
B. Short (1–59 mins)
C. Long (⩾60 mins)
Type of Performance Task
A. Reaction time
B. Attention/Perceptual
C. Mathematical processing
D. Reasoning, learning, memory
Duration of Task Battery
A. Short (<60 mins)
B. Long (⩾60 mins)

most often used as neutral conditions for comparison to hot and cold exposure. For the current meta-analysis, studies using hot temperature conditions had to use a neutral temperature condition between 60–69.9°F (15.56–21.06°C). WBGT and studies using cold temperature conditions had to use a neutral temperature condition between 65–75°F (18.33–23.89°C). Third, each study had to report on at least one type of performance measure (e.g., reaction time, tracking, memory tasks). Studies using only motor-specific tasks, self-report tasks, or physiological measures were excluded from the analysis. Last, enough reliable information had to be provided in the study to allow computation of an effect size statistic for each performance measure. In the case of the current set of data, this usually meant a direct reporting of means and standard deviations or a clear enough graph such that the data could be estimated. Based on these four criteria, 23 of the 226 primary studies could be used in the meta-analysis (Pepler 1953, Bursill 1958, Pepler 1958, Givoni and Rim 1962, Dean and McGlothlen 1964, Youngling 1965, Reilly and Parker Jr. 1968, Griffiths and Boyce 1971, Colquhoun and Goldman 1972, Reddy 1974, Bell 1978, Langkilde 1979, Epstein, *et al.* 1980, Beshir, *et al.* 1981, Lewis, *et al.* 1983, Enander 1987, Sharma and Panwar 1987, Thomas, *et al.* 1989, Armstrong and Thomas 1990, Razmjou and Kjellberg 1992, Shurtleff, *et al.* 1994, Razmjou 1996, van Orden and Benoit 1996).

It is important to note that rejecting numerous primary studies in a meta-analysis is a common occurrence. In fact, selecting only those articles that meet specific criteria is the primary means of insuring that the studies used are valid for the meta-analysis being conducted. The purpose of establishing strict criteria for inclusion is to ensure that mathematically combining across the studies that meet the criteria results in meaningful data.

2.3. *Temperature calculation*

Studies that used high temperatures as the experimental condition had to report the environmental condition either as the WBGT heat index or provide enough information to calculate the WBGT. Some studies reported the environmental temperature in the form of the Effective Temperature (ET) index. In this case, the ET was converted to WBGT using equation 1 (Brief and Confer 1971):

$$\mathrm{WBGT} = (\mathrm{ET}\text{-}13.1)/0.823 \tag{1}$$

Other studies reported the environmental temperature in the form of dry bulb (DB) and wet bulb (WB) temperatures. For these cases, it was assumed that air temperature was approximately equal to globe temperature when air movement was negligible or unreported in experiments conducted indoors (Ramsey 1995). Thus, assuming that the natural and psychrometric wet bulb temperatures were equivalent, DB and WB temperatures were converted to WBGT using equation 2 (Parsons 1995):

$$\mathrm{WBGT} = 0.7\ \mathrm{WB} + 0.3\ \mathrm{DB} \tag{2}$$

2.4. *Coding of study information*

A special coding form was developed to record pertinent information from each of the primary studies that met the criteria for inclusion. Potential moderator variables were identified by reviewing the literature related to environmental conditions and performance. The categories chosen for the current analysis are listed in the bottom portion of table 1. First, all primary studies were coded for the type of environmental temperature exposure: hot or cold. Hot temperature conditions were defined as experimental temperatures of 70°F (21.11°C) WBGT or above and were further categorized as Hot1 (70–79.9 °F [21.11–26.61°C] WBGT), Hot2 (80–89.9°F [26.67–32.17°C] WBGT), or Hot3 (≥90°F [32.22°C] WBGT). Cold temperature conditions were defined as experimental temperatures of less than 65°F (18.33°C) and were further categorized as Cold1 (50–64.9°F [10–18.28°C]) and Cold2 (<50°F [10°C]). The endpoints for the hot and cold categories were determined by the range of temperatures used in the 23 primary studies that met the decision rules for inclusion in the current meta-analysis. The endpoints for the hot and cold subcategories (Hot1, Hot2, Hot3, Cold1, Cold2) were determined after thoroughly reviewing the temperature ranges in the primary studies used in the current analysis and were chosen to equalize, as much as possible, the number of studies in each temperature subcategory. Second, the duration of the complete experimental session was coded as either short (<120 mins) or long (≥120 mins). The experimental session included any pre-task exposure to the temperature as well as the duration of all task trials and all tasks that were completed under temperature conditions. Third, the time that participants were exposed to the temperature conditions prior to working on the task (pre-task exposure) was categorized as none, short (1–59 mins),

or long ($\geqslant$60mins). Fourth, the type of task was defined as reaction time tasks, attentional or perceptual tasks (e.g., vigilance, tracking or acuity tasks), mathematical tasks (e.g., multiplication or adding tasks, identifying lower versus higher numbers), or reasoning, learning, or memory tasks (e.g., logic tasks, word recall tasks). Last, task duration was categorized as short ($<$60 mins) or long ($\geqslant$60 mins). The criteria for the potential modifier variables listed above were determined after reviewing the primary studies being used in the current analysis and examining natural cutoff points for the duration categories and logical task categories for collapsing across the primary studies. This is commonly done when choosing categories for a meta-analysis (Pilcher and Huffcutt 1996).

An effect size statistic, which indicates how many standard deviations the mean of the experimental group differed from the mean of the control group, was computed for each study using the technique described by Hunter and Schmidt (1990). The effect size statistic, *d*, was calculated using equation 3 where $\bar{X}_E$ is the mean of the experimental group, $\bar{X}_C$ is the mean of the control group, and S_p is the standard deviation pooled across both groups.

$$d = (\bar{X}_E - \bar{X}_C)/S_p \tag{3}$$

Equation 4 is the formula for computing the pooled standard deviation, where N_E and S_E represent the sample size and standard deviation for the experimental group and N_C and S_C represent the sample size and standard deviation for the control group.

$$S_p = \sqrt{[(N_E\text{-}1)S_E^2 + (N_C\text{-}1)S_C^2]/[(N_E\text{-}1) + (N_C\text{-}1)]} \tag{4}$$

In calculating effect sizes, careful attention was paid to the sign of the effect size statistic to insure that a positive *d*-score represented better performance in the experimental group than in the control group, whereas a negative *d*-score indicated worse performance.

It is important to note that most of the studies used in the current analysis used more than one of the conditions being coded, thus resulting in multiple *d*-scores for the majority of studies. For example, studies often used more than one experimental temperature condition or more than one type of performance task. A total of 517 *d*-scores were calculated from the 23 primary studies. However, the results presented here are based on 515 *d*-scores from 22 primary studies. One study (Beshir *et al.* 1981) which met the criteria for inclusion was removed from the meta-analysis following the calculation of the individual *d*-scores. The *d*-scores calculated from this study (-12.2 and -14.53) were excluded as outliers. For comparison, the *d*-scores resulting from the 22 studies used in the current meta-analysis ranged from -4.39 to 1.86.

The 22 primary studies used in the current analysis included data from 317 experimental participants. Because the data from each participant were usually included in more than one coding condition, the current results are based on a total of 7044 data points. In addition, most studies exposed their participants to both the neutral and experimental temperature conditions. Therefore, not only were multiple *d*-scores computed for many of the primary studies, the *d*-scores were not entirely independent. This commonly occurs in the meta-analytic technique. In the case of the current study, this had a minimal impact on the over-all *d*-score results. The repeated-measures within most of the studies would have underestimated the variance associated with the *d*-scores but would not have affected the *d*-scores

themselves. Thus, the pattern and strength of the *d*-scores reported here were unaffected by repeated-measures designs.

The reliability of the coding process was assessed by having two independent researchers code each of the primary studies that met the criteria for inclusion. The correlations for the *d*-scores and for each of the categories coded on the coding form were very high (ranging from 0.98 to 1.00). The few disagreements between the two raters were investigated and resolved. The high correlations between the two raters indicate that the necessary information could be coded reliably from the studies.

2.5. *Meta-analytic methodology*

All meta-analyses were completed using an SAS (SAS Institute INC, Cary, N.C.) computer program (Huffcutt, *et al.* 1992) that mathematically combines *d*-scores across primary studies. The result is an estimate of the mean effect size across the studies (i.e., the average number of standard deviations the experimental group distribution was offset from the control group distribution) and the variability observed around this average. All computations were weighted by sample size, since studies based on a larger sample are more stable than those based on a smaller sample (Hunter, *et al.* 1982, Hunter and Schmidt 1990).

It should be noted that the SAS program does not provide any tests of statistical significance. Significance testing is not typically done in the meta-analytic procedure as these procedures were developed to avoid the problems and limitations intrinsic to significance testing (Hunter, *et al.* 1982, Hunter and Schmidt 1990). Furthermore, because sampling errors tend to average out when combining across primary studies, average effect sizes from a meta-analysis represent direct estimates of the overall strength of a relationship in the population. Finally, the variability around the mean reflects the degree to which other variables moderated the relationship. Therefore, the variability does not represent a lack of consistency within the data, but, instead, provides an indication of how other variables may affect the specific variable of interest.

The overall goal of the current study was to examine the effects of hot and cold temperature exposure on performance. As a first step, a meta-analysis that collapsed across all temperature and performance task conditions was completed. This provided an overall estimate of the effect of temperature exposure on performance. Second, a separate meta-analysis for each of the major categories was completed. This included an analysis of the effect of hot and cold temperature exposure on performance, the effect of duration of temperature exposure on performance, the effect of pre-task temperature exposure on performance, the effect of the type of task on performance, and the effect of task duration on performance. The third level of the meta-analysis collapsed across the temperature subcategories and examined each of the major coding categories separately for the hot and cold temperature conditions. Finally, an additional analysis was completed for the hot temperature condition for temperatures 80°F (26.67°C) WBGT and higher (collapsing across Hot2 and Hot3 temperature subcategories).

The mean *d*-scores calculated in each of the analyses represent the mean number of standard deviations the experimental temperature group differed from the neutral temperature group for the particular variable being examined. The variance for each of the *d*-scores reflects the extent to which other experimental variables affected the magnitude of the difference between the experimental and control groups. These analyses were designed to follow a hierarchical strategy as is typically done in meta-

analytic procedures (Hunter and Schmidt 1990). In general, the current meta-analyses gradually split up the data into smaller chunks based on the coding characteristics, thus making the analyses more and more specific. By dividing the data into progressively smaller pieces, the number of studies and the number of *d*-scores being analysed at any one time became smaller and smaller. Naturally, the smaller the number of studies and d-scores in any given category, the more tentative the results.

3. Results

The results from the first stage of the analyses are presented in the top line of table 2. Combining across all coding categories, including hot and cold temperature conditions, resulted in an overall effect size of −0.192. This indicates that the mean performance of the experimental group (those participants exposed to either hot or cold temperature conditions) was 0.192 standard deviations less than the average performance of the control group (the neutral temperature group). More practically, as shown in the percentage difference column in table 2, hot and cold temperature exposure resulted in a 7.61% decrement in performance in comparison to the neutral temperature condition. Note that all percentiles presented in the tables assume that the *d*-scores approximate a normal *z* distribution and can be obtained from standard *z*-score tables.

Table 2. Meta-analysis results for each major category.

	$\bar{d}$[a]	%Diff[b]	Var($\bar{d}$)[c]	N($\bar{d}$)[d]	N(St)[e]	N(Pt)[f]
Overall	−0.192	−7.61	0.429	515	22	7044
Hot exposure	−0.150	−5.96	0.363	315	16	4229
Hot1	−0.020	−0.80	0.193	148	7	2342
Hot2	−0.189	−7.50	0.548	60	8	685
Hot3	−0.382	−14.88	0.503	107	9	1202
Cold exposure	−0.255	−10.06	0.522	200	9	2815
Cold1	−0.197	−7.81	0.381	104	4	1788
Cold2	−0.356	−13.91	0.750	96	6	1027
Short experimental session	−0.410	−15.91	0.709	189	11	1900
Long experimental session	−0.147	−5.84	0.298	264	10	4536
No pre-task temp duration	−0.101	−4.02	0.143	185	8	3277
Short pre-task temp duration	−0.127	−5.05	0.362	103	7	1090
Long pre-task temp duration	−0.472	−18.15	0.842	165	9	2069
Reaction time tasks	−0.005	−0.20	0.151	37	5	752
Attention/Perceptual tasks	−0.195	−7.73	0.379	285	16	4310
Mathematical processing tasks	−0.033	−1.32	0.259	94	7	962
Reasoning/Learning/Memory tasks	−0.469	−18.05	0.881	99	8	1020
Short task duration	−0.513	−19.60	0.850	225	12	2227
Long task duration	−0.072	−2.87	0.147	228	9	4209

[a]Mean effect size; [b]percentage difference between neutral and experimental temperature conditions; [c]variance around the mean effect size; [d]number of effect sizes; [e]number of studies; [f]number of data points in experimental groups.

The remainder of table 2 contains the results from the second stage of the analyses. As can be seen, cold exposure resulted in worse performance (10.06% decrement) than hot exposure (5.96% decrement). In both hot and cold exposure conditions, the nearer the temperature was to the neutral range the less effect it had on performance. Within hot environments, Hot1 temperatures resulted in a 0.80% decrement in performance, Hot2 temperatures resulted in a 7.50% decrement, and Hot3 temperatures resulted in a 14.88% decrement. For cold environments, Cold1 temperatures resulted in a 7.81% decrement in performance and Cold2 temperatures resulted in a 13.91% decrement.

The length of the experimental session also had an effect on performance. Short experimental sessions had a stronger negative effect on performance (15.91% decrement) than long experimental sessions (5.84% decrement). However, the longer the person was exposed to the temperature prior to task performance (pre-task temperature duration), the worse the performance. For example, long pre-task temperature duration resulted in a 18.15% decrement in performance while no pre-task temperature duration resulted in a 4.02% decrement in performance.

Performance was also affected by the type of task and task duration. Performance under environmental temperature exposure conditions was least affected for reaction time tasks (0.20% decrement) and most affected for reasoning, learning, or memory tasks (18.05% decrement). A 7.73% decrement in performance was observed in attention or perception tasks and a 1.32% decrement was observed in mathematical processing tasks. In addition, performance was more affected by short task durations (19.60% decrement) than long task durations (2.87% decrement).

The meta-analytic results from the third level of analyses examining each of the coding variables for hot and cold experimental conditions are reported in tables 3, 4, and 5. The top line in tables 3 and 4 indicate the over-all *d*-score for hot and cold

Table 3. Meta-analysis results for the hot experimental environments.

	$\bar{d}$[a]	%Diff[b]	Var($\bar{d}$)[c]	N($\bar{d}$)[d]	N(St)[e]	N(Pt)[f]
Hot exposure	−0.150	−5.96	0.363	315	16	4229
Short experimental session	−0.385	−14.99	0.512	108	6	1248
Long experimental session	−0.068	−2.71	0.264	181	9	2733
No pre-task temp duration	−0.054	−2.15	0.096	89	5	1692
Short pre-task temp duration	−0.186	−7.38	0.411	76	5	790
Long pre-task temp duration	−0.286	−11.26	0.612	124	7	1499
Reaction time tasks	−0.060	−2.39	0.241	24	4	440
Attention/Perceptual tasks	−0.194	−7.69	0.416	207	13	2885
Mathematical processing tasks	−0.098	−3.90	0.406	56	5	500
Reasoning/Learning/Memory tasks	−0.004	−0.16	0.024	28	3	404
Short task duration	−0.385	−14.99	0.670	130	6	1257
Long task duration	−0.067	−2.67	0.190	159	9	2724

[a]mean effect size; [b]percentage difference between neutral and experimental tempreature conditions; [c]variance around the mean effect size; [d]number of effect sizes; [e]number of studies; [f]number of data points in experimental groups.

Table 4. Meta-analysis results for cold experimental environments.

	$\bar{d}$[a]	%Diff[b]	Var($\bar{d}$)[c]	N($\bar{d}$)[d]	N(St)[e]	N(Pt)[f]
Cold exposure	−0.255	−10.06	0.522	200	9	2815
Short experimental session	−0.456	−17.58	1.084	81	5	652
Long experimental session	−0.268	−10.57	0.326	83	3	1803
No pre-task temp duration	−0.151	−6.0	0.188	96	5	1585
Short pre-task temp duration	0.027	1.08	0.202	27	2	300
Long pre-task temp duration	−0.963	−33.22	1.118	41	3	570
Reaction time tasks	0.072	2.87	0.013	13	2	312
Attention/Perceptual tasks	−0.197	−7.81	0.305	78	6	1425
Mathematical processing tasks	0.037	1.48	0.091	38	4	462
Reasoning/Learning/Memory tasks	−0.774	−28.05	1.209	71	6	616
Short task duration	−0.679	−25.14	1.034	95	6	970
Long task duration	−0.082	−3.27	0.068	69	2	1485

[a]mean effect size; [b]percentage difference between neutral and experimental temperature conditions; [c]variance around the mean effect size; [d]number of effect sizes; [e]number of studies; [f]number of data points in experimental groups.

Table 5. Meta-analysis results for the Hot2 and Hot3 experimental environments.

	$\bar{d}$[a]	%Diff[b]	Var($\bar{d}$)[c]	N($\bar{d}$)[d]	N(St)[e]	N(Pt)[f]
Hot2/3 exposure	−0.312	−12.25	0.528	167	12	1887
Short experimental session	−0.462	−17.80	0.504	90	6	1040
Long experimental session	−0.128	−5.09	0.498	75	6	839
No pre-task temp duration	−0.279	−10.99	0.264	19	2	338
Short pre-task temp duration	−0.252	−9.95	0.409	58	5	582
Long pre-task temp duration	−0.362	−14.13	0.691	88	6	959
Reaction time tasks	−0.064	−2.55	0.373	16	3	240
Attention/Perceptual tasks	−0.359	−14.02	0.494	121	9	1483
Mathematical processing tasks	−0.358	−13.98	1.234	26	2	120
Reasoning/Learning/Memory tasks	0.044	1.75	0.071	4	2	44
Short task duration	−0.462	−17.80	0.695	112	6	1049
Long task duration	−0.125	−4.97	0.257	53	6	830

[a]mean effect size; [b]percentage difference between neutral and experimental temperature conditions; [c]variance around the mean effect size; [d]number of effect size; [e]number of studies; [f]number of data points in experimental groups.

exposure, respectively, as reported initially in table 2. As shown in tables 3 and 4, the general pattern for hot and cold environmental temperature exposure effects on performance was similar to the over-all results, with cold exposure resulting in a greater negative effect on performance than hot exposure. However, when examining hot temperature exposure of 80°F (26.67°C) WBGT or above (table 5), hot exposure

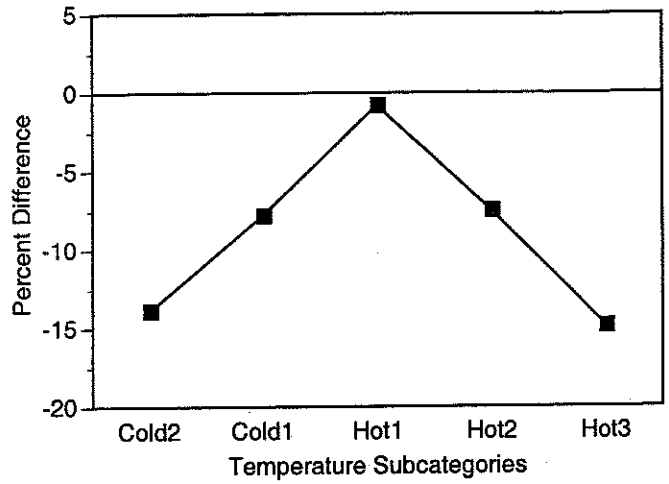

Figure 1. The mean percent difference in performance between the neutral temperature groups and the five temperature subcategories. Cold2: <50°F (10°C); Cold1: 50–64.9°F (10–18.28°C); Hot1: 70–79.9°F (21.11–26.62°C) WBGT; Hot2: 80–89.9°F (26.67–32.17°C) WBGT; Hot3: ⩾90°F (32.22°C) WBGT.

had a slightly greater negative effect on performance than cold exposure. When comparing the cold (table 4) and the Hot2 and Hot3 temperature subcategories (table 5), the general pattern remained similar to the over-all results reported in table 2. Performance was worse under short temperature exposure conditions than under long temperature exposure conditions. Long pre-task temperature durations resulted in worse performance than either no pre-task temperature duration or short pre-task temperature duration for both hot and cold. Finally, worse performance was observed in short task durations than in long task durations. The biggest difference between hot and cold temperature exposure was seen in the mathematical and the reasoning, learning and memory tasks. Performance on mathematical-related tasks was worse under hot than cold environmental conditions. In contrast, performance on reasoning, learning, and memory tasks was more negatively impacted by cold exposure than hot exposure.

Finally, to emphasize the general effect of temperature extremes on performance, figure 1 presents the mean percentage difference for each temperature subcategory (the complete data for each of the temperature subcategories are presented in table 2). As shown, the overall pattern of results is an inverted U-shape function between performance and the degree of temperature exposure. When averaging across all moderator variables within each temperature subcategory, the inverted U-shape becomes almost perfectly symmetrical between the cold subcategories and the Hot2 and Hot3 subcategories. While the Hot1 condition had very little effect on performance, the Cold2 and Hot3 temperature conditions resulted in the overall greatest decrement in performance.

4. Discussion

The current results indicate that both hot and cold temperature exposure had a negative effect on performance in a variety of different types of tasks. As expected,

the greatest detriment in performance occurred under the coldest conditions (<50 °F [10°C]) and the hottest conditions (⩾90°F [32.22°C] WBGT) – a 13.91% and a 14.88% average decrement, respectively. The Cold1 (50–64.9°F [10–18.28°C]) and Hot2 (80–89.9°F [26.67–32.17°C] WBGT) temperature conditions resulted in a smaller decrement in performance, whereas the Hot1 (70–79.9°F [21.11–26.62°C] WBGT) temperature condition had little effect on performance.

Furthermore, the effects of temperature exposure on performance varied by the type of task. Cold exposure (<65°F [18.33°C]) resulted in a large negative effect on performance on reasoning, learning, and memory tasks whereas exposure to hot environments of 80°F (26.67°C) WBGT or above on average resulted in a small improvement in performance on these types of tasks. In contrast, attentional and perceptual tasks were more negatively affected by hot exposure (⩾80°F [26.67°C] WBGT) than by cold exposure. Similarly, performance on mathematical tasks and on reaction time tasks was negatively affected by hot exposure (⩾80°F [26.67°C] WBGT) but not by cold exposure. Although these data may seem confusing when attempting to break them down as was done in tables 4 and 5, figure 1 provides an easy way to estimate the effect of temperature exposure on performance. This inverted-U shaped function may be especially useful when examining a typical work environment where the employee may be involved in a variety of tasks and a variety of environmental conditions, such as length of task duration or length of temperature exposure, simultaneously. In this type of real world setting, figure 1 provides the best estimate of the effect of temperature exposure on performance.

Although the current results do not completely agree with previous narrative summaries, there is some degree of overlap. For example, in his recent reviews, Ramsey (Ramsey and Kwon 1992, Ramsey 1995) concluded that environmental temperatures of 86 to 92°F (30–33.33°C) WBGT had a negative effect on performance of complex perceptual motor tasks but no consistent effect on mental/cognitive tasks or very simple perceptual motor tasks. With the exception of the mathematical processing tasks in the current study, these results are generally in agreement with the current results. As Ramsey did not separate mathematical processing tasks as an independent task category, it is impossible to compare the current results with Ramsey's results on those type of tasks. Another review found that mental tasks are not affected by hot exposure in the range of 86 to 92°F (30–33.33°C) WBGT but monitoring, tracking and vigilance tasks are (Echeverria, *et al.* 1991). Again, these conclusions closely match the current data with the exception of the mathematical processing tasks category. Also similar to the current results, other narrative summaries have concluded that high temperatures resulted in more severe performance decrements than more moderate temperatures (Grether 1973, Bell 1981, Hancock 1984).

In contrast to the number of reviews on the effects of hot temperature exposure, few efforts have been made to summarize the effects of cold exposure on performance. Earlier reviews examining the effects of cold exposure on performance found that more extreme cold exposure impairs performance more than moderate temperatures (Fox 1967, Hancock 1984, Enander 1987). In general, the current meta-analytic results agree with these conclusions.

One of the advantages of the current quantitative review over the previous reviews is the ability to clearly examine the effects of different moderator variables on performance. Of the variables investigated in the current study, only the duration of the experimental session has been examined in narrative reviews. In his recent

reviews, Ramsey (Ramsey and Kwon 1992, Ramsey 1995) concluded that the duration of exposure to heat was not related to performance. In contrast, the current results indicate that experimental sessions of less than 2 hours had a stronger negative impact on performance than longer durations. Similarly, tasks durations of less than 60 minutes resulted in poorer performance than longer tasks. One possible explanation for these data is that people adjust to some extent to working under extreme temperature conditions and may actually improve their performance over time when the necessary task continues throughout the temperature exposure time. This would indicate that in working environments that involve extreme temperature exposure, worse performance would be expected at the beginning of the working day than later in the day.

However, the current data also suggest that performance after a pre-task temperature duration of at least 1 hour was substantially worse than performance either immediately upon exposure to the temperature or performance within one hour of exposure to the temperature. These results would suggest that exposure to extreme temperatures prior to a specific task onset would likely bring about substantially worse performance on the new task. This would indicate that workers may need to perform the task in the temperature condition in order to adapt to the working environment and improve performance. This could be especially problematical in industrial environments where workers may be performing relatively routine tasks for a period of time under an extreme temperature condition, but after a period of exposure would not be able to completely apprehend and respond to changing parameters and performance criteria on the job.

The current results suggest that to better understand the effects of temperature exposure on performance it is necessary to better document the effects of moderator variables on performance, especially variables concerning temperature and task duration. There are also other variables that may impact performance under extreme temperature conditions such as level of acclimatization to the temperature, level of personal arousal, amount of body temperature change brought about by temperature exposure, and level of training on the task. The current study could not investigate these potential moderator variables because most studies either did not report these types of data or did not report the data in enough detail for a meta-analysis. More comprehensive meta-analyses can be completed when more studies have been published with full descriptive data.

Perhaps the major limitation of the current investigation was that many primary studies that may have been used in narrative summaries were rejected from the current meta-analysis. As explained in the methods section, rejecting many primary studies is a common occurrence in a meta-analysis. Although it is not necessary to include all possible primary articles in a meta-analytic review, it is important that a bias not be introduced inadvertently into the data. Because the studies rejected in the current meta-analysis were rejected according to rules designed to result in a scientifically sound meta-analysis, there is no a priori reason to assume that the articles rejected were systematically different in their results from the ones that were included. In addition, where they are comparable, the current overall results agree with many recent narrative summaries, indicating that the current data base was similar to the ones examined in the narrative summaries. The best way to solve this limitation is for future primary studies to include full descriptive data. This would allow for more studies to be included in a future meta-analysis and more moderator variables to be examined.

In summary, the current meta-analytic review on the effects of hot and cold temperature exposure on performance supports four major conclusions. First, the effect of cold and hot temperature exposure resulted in an inverted U-shape function with cold exposure of 50°F (10°C) or less and hot exposure of 90°F (32.22°C) WBGT or more resulting in the worse performance. In contrast, temperatures between 70 and 79.9°F (21.11–26.61°C) WBGT resulted in very little effect on performance. Second, temperature exposure had a differential effect on different types of tasks with hot exposure of over 80°F (26.67°C) WBGT having the most negative effect on attentional and perceptual type tasks and mathematical processing tasks whereas cold exposure of less than 65°F (18.33°C) had the most negative effect on reasoning, learning, and memory tasks. Third, short exposure to temperature conditions and short task durations resulted in worse performance than longer durations. Last, pre-task temperature exposure of more than 60 minutes resulted in a substantial decrement in performance. These data suggest that industries requiring workers to perform under either hot or cold temperature conditions should be aware of the potential negative effects of temperature exposure on performance.

Acknowledgements

This paper was supported by the Federal Railroad Administration's Office of Research and Development. The views of the authors do not purport to reflect the position of the Federal Railroad Administration or the Department of Transportation. We thank Dr. Thomas Raslear for his valuable comments on the data and the manuscript.

References

Armstrong, D. W. and Thomas, J. R. 1990, *Effects of multiple cold air exposures on delayed matching to sample performance* (Washington, DC: Naval Medical Command), NMRI 90–87.

Azer, N. Z., McNall, P. E. and Leung, H. C. 1972, Effects of heat stress on performance, *Ergonomics*, **15**, 681–691.

Bell, C. R. and Provins, K. A. 1962, Effects of high temperature environmental conditions on human performance, *Occupational Medicine*, **4**, 202–211.

Bell, C. R., Provins, K. A. and Hiorns, R. W. 1964, Visual and auditory vigilance during exposure to hot and humid conditions, *Ergonomics*, **7**, 279–288.

Bell, P. A. 1978, Effects of noise and heat stress on primary and subsidiary task performance, *Human Factors*, **20**, 749–752.

Bell, P. A. 1981, Physiological, comfort, performance, and social effects of heat stress, *Journal of Social Issues*, **37**, 71–94.

Beshir, M. Y., El-Sabagh, A. S. and El-Nawawi, M. A. 1981, Time on task effect on tracking performance under heat stress, *Ergonomics*, **24**, 95–102.

Brief, R. S. and Confer, R. G. 1971, Comparison of heat stress indices, *American Industrial Hygiene Association Journal*, **32**, 11–16.

Bursill, A. E. 1958, The restriction of peripheral vision during exposure to hot and humid conditions, *The Quarterly Journal of Experimental Psychology*, **10**, 113–129.

Chiles, W. D. 1958, Effects of elevated temperatures on performance of a complex mental task, *Ergonomics*, **2**, 89–96.

Colquhoun, W. P. 1969, Effects of raised ambient temperature and event rate on vigilance performance, *Aerospace Medicine*, **40**, 413–417.

Colquhoun, W. P. and Goldman, R. F. 1972, Vigilance under induced hyperthermia, *Ergonomics*, **15**, 621–632.

DEAN, R. D. and MCGLOTHLEN, C. L. 1964, Effects of combined heat and noise on human performance, physiology, and subjective estimates of comfort and performance, In *Proceedings of the Institute of Environmental Science Annual Technical Meeting*, 55–64.

ECHEVERRIA, D., BARNES, V. and BITTNER JR., A. 1991, The impact of environmental exposures on industrial performance tasks. In W. Karwowski and J. W. Yates (eds) *Advances in Industrial Ergonomics and Safety III*, (London: Taylor & Francis), 629–636.

ENANDER, A. 1987, Effects of moderate cold on performance of psychomotor and cognitive tasks, *Ergonomics*, **30,** 1431–1445.

EPSTEIN, Y., KEREN, G., MOISSEIEV, J., GASKO, O. and YACHIN, S. 1980, Psychomotor deterioration during exposure to heat, *Aviation, Space, and Environmental Medicine*, **51,** 607–610.

FINE, B. J. and KOBRICK, J. L. 1978, Effects of altitude and heat on complex cognitive tasks, *Human Factors*, **20,** 115–122.

FOX, W. F. 1967, Human performance in the cold, *Human Factors*, **9,** 203–220.

FRASER, D. C. 1957, Environmental stress and its effects on performance, *Occupational Psychology*, **21**, 248–255.

GIVONI, B. and RIM, Y. 1962, Effect of the thermal environment and psychological factors upon subjects' responses and performance of mental work, *Ergonomics*, **5,** 99–114.

GOYDER, J. and MCCUTCHEON, T. I. 1995, Francophone life satisfaction and civic culture: A meta-analysis of the Canadian case, *Social Indicators Research*, **34,** 377–394.

GRETHER, W. F. 1973, Human performance at elevated environmental temperatures, *Aerospace Medicine*, **44,** 747–755.

GRETHER, W. F., HARRIS, C. S., MOHR, G. C., NIXON, C. W., OHLBAUM, M., SOMMER, H. C., THALER, V. H. and VEGHTE, J. H. 1971, Effects of combined heat, noise and vibration stress on human performance and physiological functions, *Aerospace Medicine*, **42,** 1092–1097.

GRIFFITHS, I. D. and BOYCE, P. R. 1971, Performance and thermal comfort, *Ergonomics*, **14,** 457–468.

HANCOCK, P. A. 1984, Environmental stressors. In J. S. Warm (ed.) *Sustained Attention n Human Performance*, (Chichester: Wiley), 103–142.

HUFFCUTT, A., ROTH, P. and MCDANIEL, M. 1996, A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity, *Journal of Applied Psychology*, **81**, 459–473.

HUFFCUTT, A. I., ARTHUR, W. J. and BENNETT, W. 1992, Conducting meta-analysis using the 'PROC MEANS' procedure in SAS, *Educational and Psychological Measurement*, **53,** 119–131.

HUNTER, J. E. and SCHMIDT, F. L. 1990, *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. (Newbury Park, CA: Sage Publications).

HUNTER, J. E., SCHMIDT, F. L. and JACKSON, G. B. 1982, *Meta-Analysis: Cumulating Research Findings Across Studies*. (Beverly Hills, CA: Sage Publications).

HYMAN, R. B., FELDMAN, H. R., HARRIS, R. B., LEVIN, R. F. and MALLOY, G. B. 1989, The effects of relaxation training on clinical symptoms: A meta-analysis, *Nursing Research*, **38,** 216–220.

KLAWANSKY, S., YEUNG, A., BERKEY, C., SHAH, N., PHAN, H. and CHALMERS, T. C. 1995, Meta-analysis of randomized controlled trials of cranial electrostimulation: efficacy in treating selected psychological and physiological conditions, *The Journal of Nervous and Mental Disease*, **183,** 478–484.

KOBRICK, J. L. and FINE, B. J. 1983, Climate and human performance. In D. J. Oborne and M. M. Gruneberg (eds) *The Physical Environment at Work*, (Chichester: Wiley), 69–107.

KOBRICK, J. L. and JOHNSON, R. F. 1991, Effects of hot and cold environments on military performance. In R. Gal and A. D. Mangelsdorff (eds) *Handbook of Military Psychology*, (Chichester: Wiley), 215–232.

LANGKILDE, G. 1979, The influence of thermal environment on office work. In P. O. Fanger and O. Valbjorn (eds) *Indoor Climate: Effects of human comfort, performance, and health in esidential, commercial, and light-industry buildings* (Copenhagen: Danish Building Research Institute), 835–856.

LEWIS, M. I., MEESE, G. B. and KOK, R. 1983, *The effects of moderate cold and heat stress on the potential work performance of industrial workers* (Pretoria: National Building Research, Institute Council for Scientific and Industrial Research), CSIR Rep. No. 589.

MACKWORTH, N. H. 1947, The breakdown of vigilance during prolonged visual search, *Quarterly Journal of Experimental Psychology*, 6–21.

NIOSH 1972, *Criteria for a Recommended Standard-Occupational Exposure to Hot Environments* (Washington, DC: National Institute for Occupational Safety and Health), HSM 72-10269.

NIOSH 1980, *Proceedings of a Workshop on Recommended Heat Stress Standards*, F. N. Dukes-Dobos and A. Henschel (eds) (Cincinnati, OH: National Institute of Occupational Safety and Health), DDHS (NIOSH) 81–108.

NIOSH 1986, *Criteria for Recommended Standard-Occupational Exposure to Hot Environments Revised Criteria 1986* (Washington, DC: National Institute for Occupational Safety and Health), DDHS (NIOSH) 86–113.

PARSONS, K. C. 1995, International heat stress standards: A review, *Ergonomics*, **38,** 6–22.

PAUS, T., KOSKI, L., CARAMANOS, Z. and WESTBURY, C. 1998, Regional differences in the effects of task difficulty and motor output on bloodflow response in the human anterior cingulate cortex: A review of 107 PET activation studies, *NeuroReport*, **9,** R37–R47.

PEERS, I. S. and JOHNSON, M. 1994, Influence of learning context on the relationship between A-level attainment and final degree performance: A meta-analytic review, *British Journal of Educational Psychology*, **64,** 1–18.

PEPLER, R. D. 1953, *The effect of climatic factors on the performance of skilled tasks by young European men living in the tropics* (London: MRC, Royal Naval Tropical Research Unit), Rep. No. 22161.

PEPLER, R. D. 1958, Warmth and performance: An investigation in the tropics, *Ergonomics*, **2,** 63–88.

PEPLER, R. D. 1959, Warmth and lack of sleep: Accuracy or activity reduced? *Journal of Comparative and Physiological Psychology*, **52,** 446–475.

PEPLER, R. D. 1960, Warmth, glare and a background of quiet speech: A comparison of their effects on performance, *Ergonomics*, **3,** 68–73.

PILCHER, J. J. and HUFFCUTT, A. I. 1996, Effects of sleep deprivation on performance: A meta-analysis, *Sleep*, **19,** 318–326.

PILCHER, J. J., LAMBERT, B. J. and HUFFCUTT, A. I. 2000, Differential effects of permanent and rotating shifts on self-report sleep length: A meta-analytic review, *Sleep*, **23,** 155–163.

RAMSEY, J. D., DAYAL, D. and GHAHRAMANI, B. 1975, Heat stress limits for sedentary workers, *American Industrial Hygiene Association Journal*, **36,** 259–265.

RAMSEY, J. D. 1983, Heat and cold. In G. R. J. Hockey (ed.) *Stress and Fatigue in Human Performance* (Chichester: Wiley), 33–60.

RAMSEY, J. D. 1995, Task performance in heat: A review, *Ergonomics*, **38,** 154–165.

RAMSEY, J. D. and KWON, Y. G. 1992, Recommended alert limits for perceptual motor loss in hot environments, *International Journal of Industrial Ergonomics*, **9,** 245–257.

RAMSEY, J. D. and MORRISSEY, S. J. 1978, Isodecrement curves for task performance in hot environments, *Applied Ergonomics*, **9,** 66–72.

RAMSEY, J. D. and PAI, S. B. 1975, Sedentary work by females in hot environments. In *Proceedings of the Human Factors Society 19th Annual Meeting* (Santa Monica, CA: Human Factors Society), 498–500.

RAZMJOU, S. 1996, Mental workload in heat: Toward a framework for analyses of stress state, *Aviation, Space, and Environmental Medicine*, **67,** 530–538.

RAZMJOU, S. and KJELLBERG, A. 1992, Sustained attention and serial responding in heat: Mental effort in the control of performance, *Aviation, Space, and Environmental Medicine*, **63,** 594–601.

REDDY, S. P. 1974, Sedentary job performance within the thermal comfort zone. Unpublished master's thesis, Texas Tech University.

REILLY, R. E. and PARKER JR., J. F. 1968, *Effect of heat stress and prolonged activity on perceptual-motor performance* (Arlington, VA: Biotechnology, Inc), NASA Contractors Rep. CR-1153.

SHARMA, V. M. and PANWAR, M. R. 1987, Variations in mental performances under moderate cold stress, *International Journal of Biometeorology*, **31,** 85–91.

SHURTLEFF, D., THOMAS, J. R., SCHROT, J., KOWALSKI, K. and HARFORD, R. 1994, Tyrosine reverses a cold-induced working memory deficit in humans, *Pharmacology, Biochemistry, and Behavior*, **47,** 935–941.

SVARTBERG, M. and STILES, T. C. 1991, Comparative effects of short-term psychodynamic psychotherapy: A meta-analysis, *Journal of Consulting and Clinical Psychology*, **59,** 704–714.

THOMAS, J. R., AHLERS, S. T., HOUSE, J. F. and SCHROT, J. 1989, Repeated exposure to moderate cold impairs matching-to-sample performance, *Aviation, Space, and Environmental Medicine*, **60,** 1063–1067.

VAN ORDEN, K. F. and BENOIT, S. L. 1996, Effects of cold air stress on the performance of a command and control task, *Human Factors*, **38,** 130–141.

WILKINSON, R. 1969, Some factors influencing the effect of environmental stressors upon performance, *Psychological Bulletin*, **72,** 260–272.

YOUNGLING, E. W. 1965, The effects of thermal environment s and sleep deprivation upon concurrent central and peripheral tasks. Unpublished doctoral dissertation, University of Massachusetts, Amherst.